Homework 2
Due: Wednesday, Oct 6, 11:59pm EST on Gradescope

Exercises (Please Solve but Do Not Submit)

1. Bourgain's Embedding. Some questions about Bourgain's embeddings.

(a) Prove that for any p, the same random-subsets embedding given in Lecture 7 gives an $O(\log n)$-distortion embedding into ℓ_p, with high probability. (Hint: Use Hölder's inequality, that for any two vectors a, b, and any $p, q \geq 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, we have $\langle a, b\rangle \leq \|a\|_p \|b\|_q$.)

(b) Given a metric (V, d), and a subset $T \subseteq V$, alter Bourgain's construction to construct a map $\varphi : V \to \mathbb{R}^{O(\log n \log |T|)}$ such that (a) $\|\varphi(i) - \varphi(j)\| \leq d(i, j)$ for all $i, j \in V$, but (b) $\|\varphi(i) - \varphi(j)\| \geq \frac{d(i,j)}{\alpha}$ for all $i, j \in T$, such that $\alpha = O(\log |T|)$, with high probability. Show that it implies an $O(\log |T|)$-approximation to Generalized Sparsest Cut.

2. Dimension of a Set System. Given a set system $(U, \mathcal{S})$, say a set $A \subseteq U$ can be *shattered* if for each subset A' of A, there is some set $S \in \mathcal{S}$ such that $A' = S \cap A$. In other words,

$$|\{S \cap A \mid S \in \mathcal{S}\}| = 2^{|A|}.$$

The *Vapnik-Chervonenkis (VC) dimension* of $(U, \mathcal{S})$ is the largest size of a set $A \subseteq U$ that can be shattered by $\mathcal{S}$.

(a) Let $U = \mathbb{R}^2$ and $\mathcal{S}$ contain all half-spaces. Show that no set of size 4 can be shattered by $\mathcal{S}$. Hence infer that the VC dimension of this set system is 3. (In fact, the VC dimension of half-spaces in $\mathbb{R}^d$ can be shown to be $d + 1$.)

(b) What is the VC dimension of the set of intervals in $\mathbb{R}$?

Suppose there is a probability distribution μ over U, such that $\mu(S)$ is well-defined for every $S \in \mathcal{S}$ (i.e., these sets are "measurable"). A set $N \subseteq U$ is called an *ε-net* for the set system if for every set $S \in \mathcal{S}$ with $\mu(S) \geq \varepsilon$ (i.e., the set is "large"), we have $S \cap N \neq \varnothing$.

(c) Show that a random sample of U of $O(\frac{1}{\varepsilon} \log |\mathcal{S}|)$ points from the distribution μ is an ε-net with constant probability.

Such a result is not interesting when $\mathcal{S}$ is large. A surprising theorem of Haussler says that for any set system of VC dimension d, there exists an ε-net of size $O(\frac{d}{\varepsilon} \log \frac{1}{\varepsilon})$—in fact, that a random sample of U of this size is an ε-net with constant probability. (The proof is somewhat tricky, and we will discuss it some other time.)

(d) Sometimes one gets better ε-nets than the above construction. Show there exists ε-net for intervals on the line of size $O(1/\varepsilon)$.

3. Max-Flow/Min-Cut using Duality and LP Rounding. Given a directed graph $G = (V, E)$ with vertices s, t and edge capacities c_e. Consider the following pair of LPs, where the primal variables f_P are flows on s-t paths P, and the dual variables are lengths y_e on edges $e \in E$.

(P) : max	$\sum_P f_p$	(D) : min	$\sum_e c_e y_e$
	$\sum_{P:e\in P} f_P \le c_e$		$\sum_{e\in P} y_e \ge 1$
	$f_P \ge 0.$		$y \ge 0.$

(The primal has exponentially many variables, the dual many constraints, but let's not worry about that.) These are dual LPs, and hence by strong LP duality they have equal values (assuming both are feasible. The primal is a formulation of max-flow. The dual seems like *a relaxation* of the min-s-t-cut problem: while setting $y_e = 1$ on the edges of any s-t cut is an integer solution, we also allow fractional solutions. However, let's show this LP has no integrality gap, so there always exist integer optimal solutions.

(i) For each vertex v, let d_v be the shortest-path distance from s to v, according to edge lengths $y_e \ge 0$. Choose $\alpha \in [0, 1)$ uniformly at random. Define $S_\alpha := \{v \in V \mid d_v \le \alpha\}$. Show that

$$\mathbb{E}_\alpha[c(S_\alpha, \bar{S}_\alpha)] := \mathbb{E}_\alpha\big[\textstyle\sum_{e\in S_\alpha} c_e\big] \le \sum_e c_e y_e.$$

(ii) Infer that $\min_\alpha c(S_\alpha, \bar{S}_\alpha)$ is a min-s-t-cut. (Make sure you see why we could not choose $\alpha \in [0, 2]$?) Given $\{y_e\}_{e\in E}$ find this min-cut in poly-time. (We will see how to solve the LP in poly-time as well, later in the course.)

This proves max-flow = min-cut, albeit using the heavier hammer of strong LP duality.

4. SDP Rounding for Max 2-XOR In this problem, we will analyze an algorithm for a generalization of the Max Cut problem. In the Max 2-XOR problem, the input is an undirected graph $G(V, E)$ on $n = |V|$ vertices and $m = |E|$ edges along with "right-hand sides" $b_{i,j} \in \{-1, 1\}$, one for each $\{i, j\} \in E$. We say that an $x \in \{-1, 1\}^n$ *satisfies* an edge $\{i, j\} \in E$ if $x_i x_j = b_{i,j}$. The 2-XOR value of the instance is the maximum over all $x \in \{-1, 1\}^n$ of the fraction of the edges satisfied by x. The Max 2-XOR problem is to take input $G, \{b_{i,j}\}_{\{i,j\}\in E}$ and output an x that satisfies the maximum possible fraction of edges.

(i) Convince yourself that if $b_{i,j} = 1$ for each $\{i, j\} \in E$, then, Max 2-XOR is same as the Max-Cut problem on graph G.

(ii) Let $A \in \mathbb{R}^{n\times n}$ be the matrix such that $A_{i,j} = A_{j,i} = b_{i,j}$ if $\{i, j\} \in E$ and $A_{i,j} = A_{j,i} = 0$ otherwise. Consider the semidefinite program P:

$$\max \frac{1}{2} + \frac{1}{4m}\langle A, X\rangle \text{ s.t.} \{X \in \mathbb{R}^{n\times n} \mid X \succeq 0,\ X_{i,i} = 1 \text{ for each } 1 \le i \le n\}.$$

(Recall from HW#0: $\langle A, X\rangle = \sum_{i,j} A_{i,j} X_{i,j}$ is the Frobenius inner product on matrices.) Prove that P is a relaxation of the Max 2-XOR problem.

(iii) Let's analyze the Goemans-Williamson rounding algorithm for this problem. Let $X = VV^\top$ be a Cholesky decomposition of X for matrix $V \in \mathbb{R}^{n\times n}$ with rows $V_1, V_2, \ldots, V_n$. Let $g \sim \mathbb{N}(0, 1)^n$ be a standard Gaussian vector (i.e., each g_i is an independent draw from $\mathcal{N}(0, 1)$). For each i, set $x_i = \text{sign}(\langle g, V_i\rangle)$.

Use the analysis from class to prove that if the value of the input instance is $1 - \epsilon$ for any $\epsilon > 0$, then $1 - O(\sqrt{\epsilon})$ fraction of constraints are satisfied by x in expectation.

Problems

1. Adding to a Tree. In the TREE ADDITION problem, you are given a tree $T = (V, E)$ and a collection of edges $J = \{e_i = \{u_i, v_i\}\}_{i=1}^m$ (called *jumps*) on the same set V of vertices. Adding a jump to a tree creates a cycle, and we say that the jump "covers" the tree edges on this cycle. (In other words, all edges on the unique path between vertices u_i and v_i are covered by jump e_i.) The jumps have costs, and we want to add the smallest-cost set of jumps that cover all tree edges.

(a) (Do not submit.) Given an $O(\log n)$ approximation, where $n = |V|$.

(b) Suppose T has a root node r: this defines a notion of ancestors/descendents. Suppose all the jumps e_i have the property that one end (say u_i) is a descendent of the other end v_i in T. Solve this problem in polynomial time. (Hint: dynamic programming.) If you are stuck, start with the case where T is a path, and the root is one end of this path.

(c) Use the previous part to give a 2-approximation for the general case, where jumps can go between nodes that are not ancestors/descendents.

2. Maximizing Quadratics. We now design an algorithm that, given a homogenous degree-2 polynomial $p(x) = \sum_{i,j} p_{i,j} x_i x_j$, outputs an approximate maximizer of p over $x \in \{-1, 1\}^n$. We will assume, in addition, that $p_{i,i} = 0$ for each $1 \leq i \leq n$.

(i) **Basic Algorithm.** (2 points) Let's give a baseline solution which will be useful in the final analysis of the rounding. Prove that $\max_{x \in \{-1,1\}^n} p(x) \geq \frac{1}{n} \sum_{i,j} |p_{i,j}|$.

(Hint: Let M be a perfect matching of the complete graph on $[n]$. For each edge $\{i, j\} \in M$, let Y_i be set uniformly at random from $\{-1, 1\}$ and let $Y_j = Y_i$ if $p_{i,j} > 0$ and $Y_j = -Y_i$ otherwise. Prove that $\mathbb{E}[p(Y)] = \sum_{\{i,j\} \in M} |p_{i,j}|$. Now think about choosing M randomly and repeating the above analysis.)

(ii) **Interval constraints do not help.** (1 points) Prove that for a polynomial p as above, the maximum of p over $x \in [-1, 1]^n$ is equal to the maximum of p over $x \in \{-1, 1\}^n$. *(Hint: Think of a natural "randomized" rounding of a maximizer $y \in [-1, 1]^n$ into a $x \in \{-1, 1\}^n$ such that $\mathbb{E}[p(x)] = p(y)$.)*

(iii) **SDP Relaxation.** (1 points) Let $A \in \mathbb{R}^{n \times n}$ be the matrix such that $A_{i,j} = A_{j,i} = p_{i,j}$ for each $\{i, j\} \in E$ and $A_{i,j} = A_{j,i} = 0$ otherwise. Consider the semidefinite program:

$$\max\{\langle A, X \rangle \mid X \in \mathbb{R}^{n \times n}, X \succeq 0, \ X_{i,i} = 1 \text{ for each } 1 \leq i \leq n\}. \tag{1}$$

Show this is a relaxation for the problem.

We will now give a rounding algorithm for the SDP. Given part (ii) it suffices to give a randomized rounding algorithm that takes X and outputs a fractional solution $x \in [-1, 1]^n$. Let X be an optimal solution to (1) and let OPT_{SDP} be the optimum value. Let $X = VV^\top$ be a Cholesky decomposition of X for some matrix $V \in \mathbb{R}^{n \times n}$ with rows $V_1, V_2, \ldots, V_n \in \mathbb{R}^n$. Let $g \sim \mathcal{N}(0, 1)^n$ be a standard Gaussian vector. For some $B > 0$ (to be chosen later), let $z_i = \frac{1}{B} \langle g, V_i \rangle$ for each $1 \leq i \leq n$. If $|z_i| \leq 1$, set $x_i = z_i$. Otherwise, set $x_i = \text{sign}(z_i)$.

In the following parts, we analyze this rounding algorithm.

(iv) Prove that for every i, j, $\mathbb{E}[z_i z_j] = \frac{1}{B^2} X_{i,j}$.

(v) Let R be the event that $\{|z_i| \leq 1 \text{ and } |z_j| \leq 1\}$. Prove that $\mathbb{E}[(z_i z_j - x_i x_j)\mathbf{1}_R] = 0$, where $\mathbf{1}_R$ indicates the 0-1 indicator for event R.

(vi) Let $\bar{R}$ be the complement of the event R. Prove that $|\mathbb{E}[x_i x_j \mathbf{1}_{\bar{R}}]| \leq 2e^{-B^2/2}$. (Hint: give an upper bound on the chance that $\bar{R}$ happens.)

(vii) Prove that for any i, j, $|\mathbb{E}[z_i z_j \mathbf{1}_{\bar{R}}]| \leq 100e^{-B^2/2}$.

(viii) Use bounds from previous 3 parts to conclude that $|\mathbb{E}[z_i z_j] - \mathbb{E}[x_i x_j]| \leq 1000e^{-B^2/2}$. Derive that $\mathbb{E}[p(x)] \geq \frac{1}{B^2} OPT_{SDP} - 1000\, e^{-B^2/2} \sum_{i,j} |p_{i,j}|$.

(ix) Set $B = O(\sqrt{\log n})$. Use the basic algorithm along with the conclusion of (viii) to conclude that the algorithm provides an approximation ratio of $O(\log n)$ for our problem.

Some facts you may use without proof. If $z \sim \mathcal{N}(0,1)$, for any $t \geq 1$, the following hold (1) $\Pr[z > t] \leq e^{-t^2/2}$, (2) $\mathbb{E}[|z| \cdot \mathbf{1}_{(z \geq t)}] \leq e^{-t^2/2}$, (3) $\mathbb{E}[z^2 \cdot \mathbf{1}_{(z \geq t)}] \leq 10te^{-t^2/2}$.

3. Low-Dimensional Hitting Set. Suppose $(U, \mathcal{S})$ is a set system with VC dimension d. (See the exercises for definitions.) Recall the hitting set problem from HW#1 exercises, which you argued was equivalent to the set cover problem. Let x^* be an optimal solution to the hitting set LP

$$\min \left\{ \sum_e x_e \mid \sum_{e \in S} x_e \geq 1 \; \forall S \in \mathcal{S},\; x_e \geq 0 \right\},$$

and let $LP^* = \sum_e x_e^*$ be the optimal LP value. Give an $O(d \log LP^*)$-approximation for (unit weight) hitting set on this set system. *Hint: define a probability distribution μ over U, such that $\mu(S) \geq 1/LP^*$ for all sets $S \in \mathcal{S}$.*

4. Graph Coloring and Eigenvalues Let $G = (V, E)$ be a graph with n vertices and $nd/2$ edges. Let $d_{\max}$ be the maximum degree of any vertex in G. A proper *k-coloring* of G is a map of vertices to colors from $\{1, 2, \ldots, k\}$ so that every edge has endpoints with different colors.

(i) (1 point) Prove that G has a proper coloring using $d_{\max} + 1$ colors. (Hint: be greedy!)

(ii) (2 points) Suppose that for every i, vertex i has at most r neighbors in $\{1, 2, \ldots, i-1\}$. Prove that G has a proper coloring with $r + 1$ colors.

(iii) (2 points) Prove that $d_{\max} \geq \lambda_{\max}(A) \geq d$ where $\lambda_{\max}(A)$ denotes the largest eigenvalue of the adjacency matrix A of G. You may want to use the characterization and bounds on $\lambda_{\max}$ from the exercises in HW0.

(iv) (2 points) Let B be a matrix obtained by removing the i^{th} row and i^{th} column of A for any i. Prove that the $\lambda_{\max}(B) \leq \lambda_{\max}(A)$.

(v) (3 points) Infer that there is an ordering of vertices of G such that every vertex v has at most $\lfloor \lambda_{\max}(A) \rfloor$ neighbors in G that appear before v in the ordering.

Hence, you get that G has a proper k-coloring for $k = \lfloor \lambda_{\max}(A) \rfloor + 1$. This is never worse than the bound in (i) above, and may be much smaller.